USE OF BIOINFORMATIC APPROACHES TO UNAMBIOGUOUSLY INDENTIFY HOST GENE RESPONSES CHARACTERISTIC OF EXPOSURE TO BIOTHREAT AGENTS

Rasha Hammamieh and Marti Jett

Department of Molecular Pathology

Walter Reed Army Institute of Research

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
18 NOV 2003	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Use Of Bioinformatic Approaches To Unambioguously Indentify Host Gene Responses Characteristic Of Exposure To Biothreat Agents	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Department of Molecular Pathology Walter Reed Army Institute of Research**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM001851, Proceedings of the 2003 Joint Service Scientific Conference on Chemical & Biological Defense Research, 17-20 November 2003.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**34**	

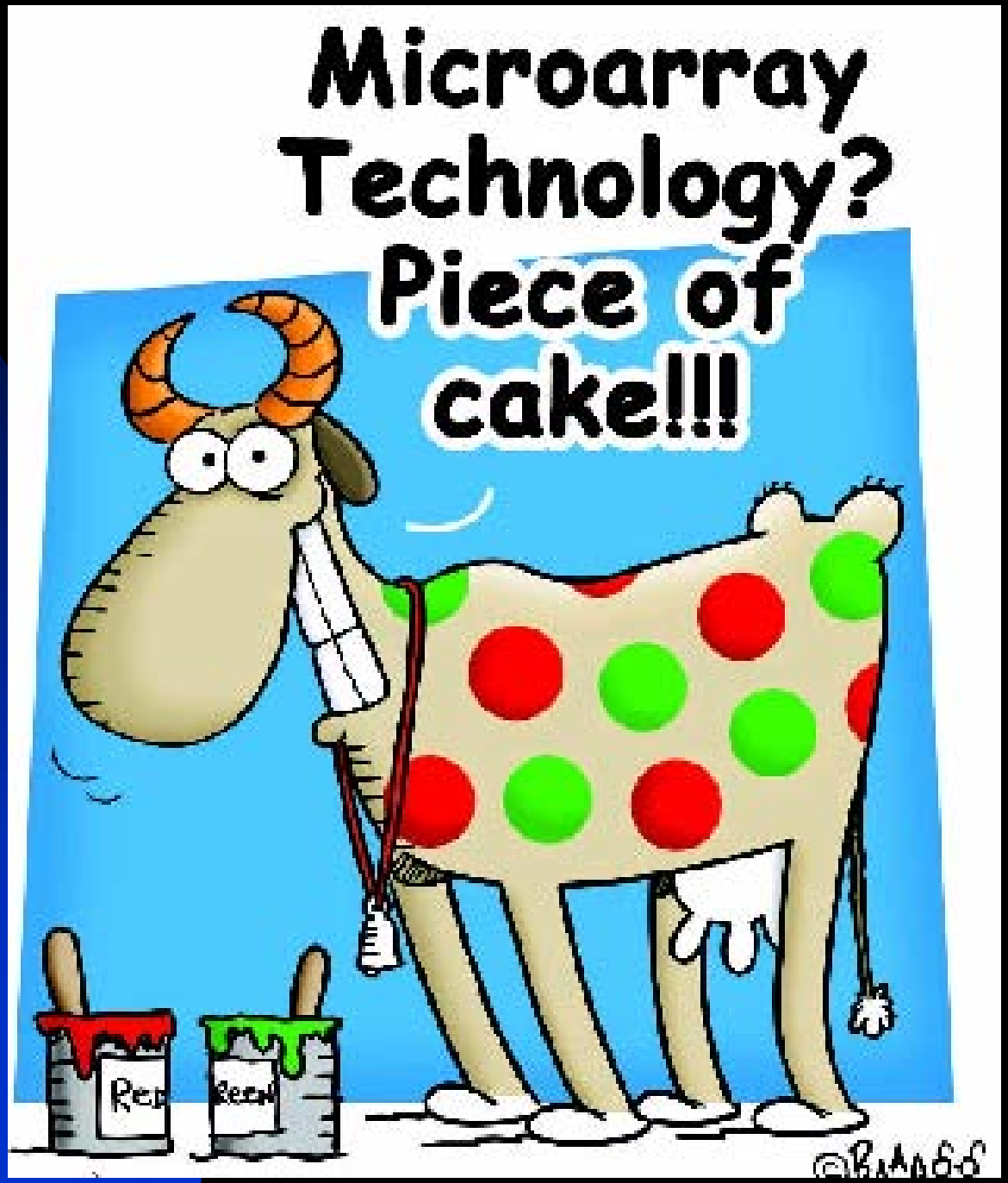
Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

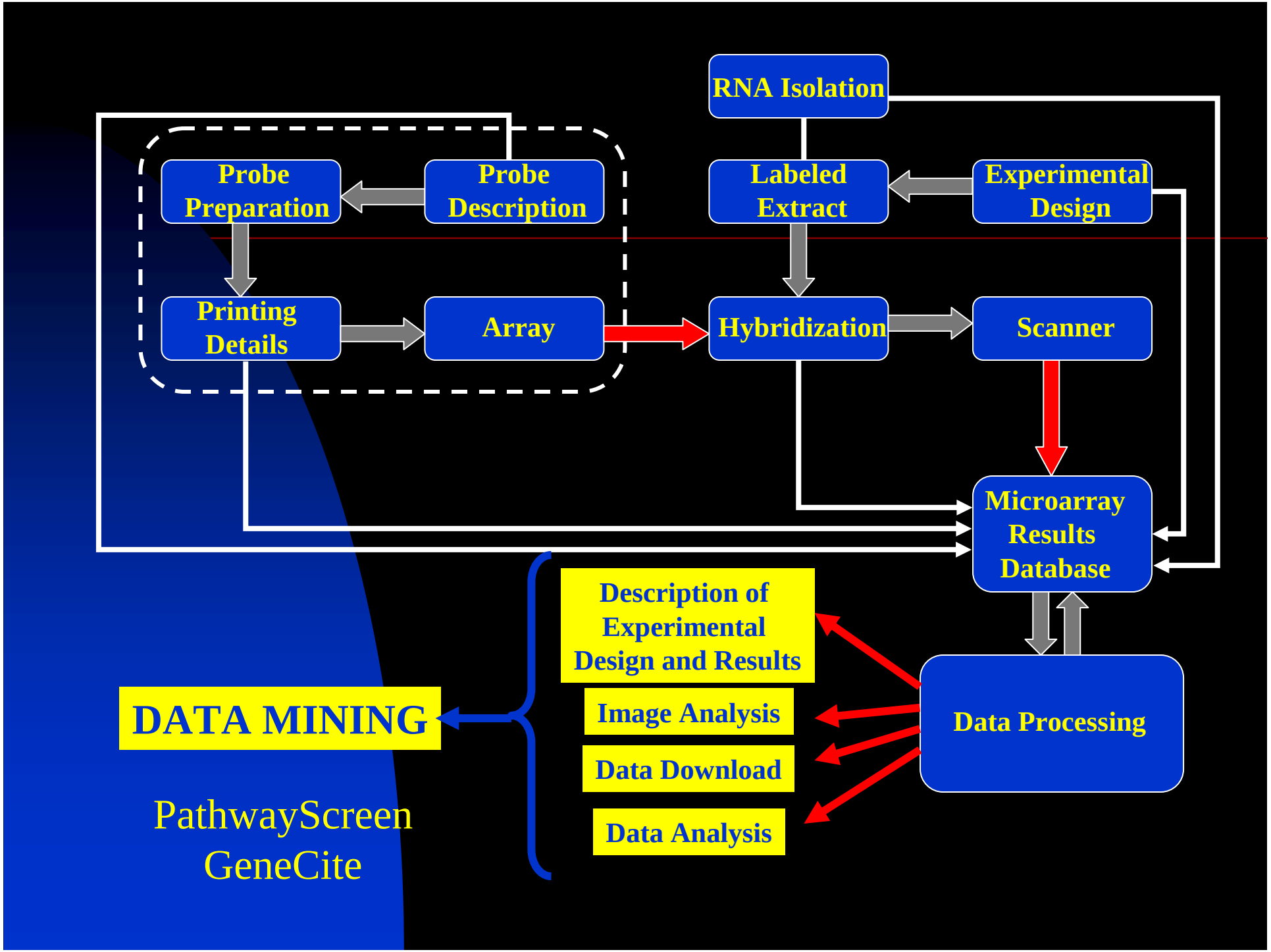
- Overview
- Establishment of base line gene expression in the host PBMC
  - Project Normal
- Data mining and feature selection
  - Genes as diagnostic markers for various pathogens.
  - *In vitro* vs. *in vivo*.
  - Applications.

Department of Molecular Pathology: Dr. Jett
Anthrax (G. Ludwig)
VEE(G. Ludwig)
CT
SEs (SEA, SEB, SEC, TSST…)
Plague (L. Lindler)
Brucella (D. Hoover)
Botulinum (M. Jensen and L. Smith)
Dengue (W. Sun)
Influenza A, B and Parainfluenzas I, III

Using:
DD-PCR
Real Time Rt-PCR
cDNA arrays
Customized cDNA glass chips

Systems Used:
In vitro:
Human peripheral blood mononuclear cells exposed to Biological threat agents.
In vivo:
Monkeys exposed to Biological threat agents (Kotersky, Ludwig)
Piglets exposed to SEs (Mani and Bi)


Microarray Technology?
Piece of cake!!!

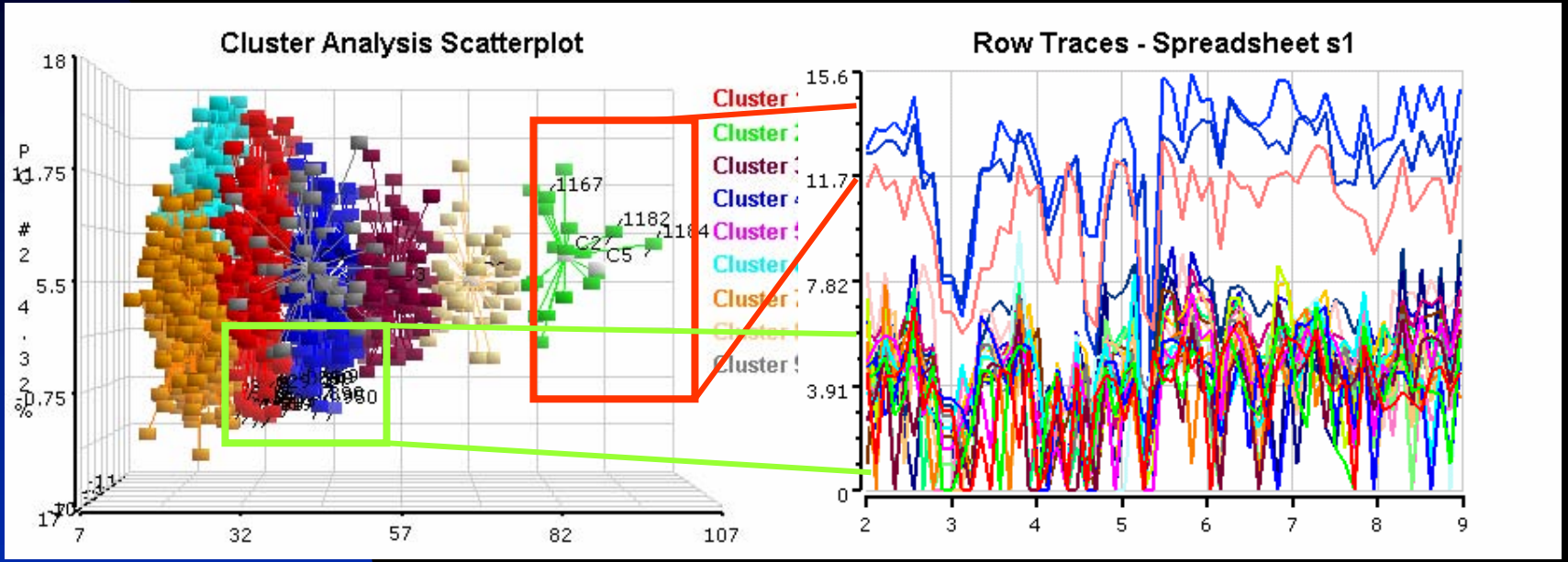

RNA Isolation
Probe Preparation
Probe Description
Labeled Extract
Experimental Design
Printing Details
Array
Hybridization
Scanner
Microarray Results Database
Description of Experimental Design and Results
Image Analysis
Data Download
Data Analysis
Data Processing
DATA MINING
PathwayScreen
GeneCite

- Overview
- Establishment of base line gene expression in the host PBMC
  - Project Normal
- Data mining and feature selection
  - Genes as diagnostic markers for various pathogens.
  - *In vitro* vs. *in vivo*.
  - Applications.

Project Normal: genetic variation and gene expression baseline for human PBMC

- When assaying the expression of thousands of transcripts using microarrays, there is a high likelihood of finding differentially expressed genes that vary normally in the studied tissues.
- Before interpreting microarray data, it is necessary to define the normal physiological variance in gene expression.

  Samples were collected from 75 donors, healthy, ethnically diverse, gender and age.
- Using cDNA microarray we examined the variance in transcript levels for thousands of genes in normal human PBMC.


Cluster Analysis Scatterplot
Row Traces - Spreadsheet s1
18
11.75
5.5
0.75
7
32
57
82
107
1167
1182
C2
C5
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
15.6
11.7
7.82
3.91
0
2
3
4
5
6
7
8
9

Project Normal: Normally Variable Genes.

C:/DOCUME~1/HAMMAM~1/LOCALS~1/Temp/64

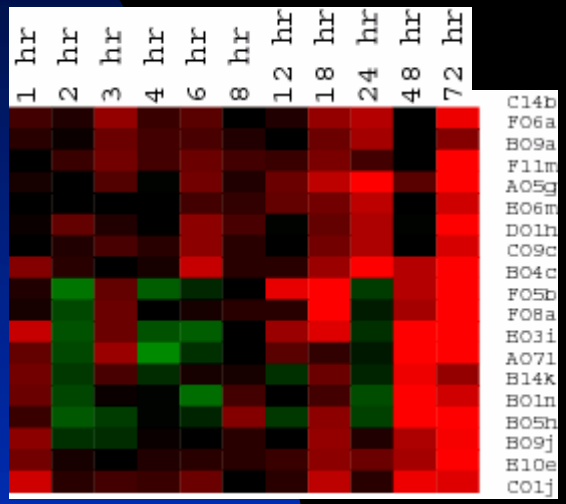
Column	Column ID	p-value	p-value (corrected for multiple comparisons)	Model DOF	Error DOF	Model Sum of Squares
771	MCM5 DNA replication licensing factor; CDC46 homolog	1.36E-08	1.61E-05	7	55	5.0795
1145	interleukin-7 receptor alpha subunit precursor (IL-7R-alpha; IL7R); CDW127 antigen	2.18E-08	2.58E-05	7	55	10.9483
793	DNA excision repair protein ERCC1	3.22E-08	3.82E-05	7	55	5.24571
1026	histidine decarboxylase (HDC)	4.85E-08	5.74E-05	7	55	9.30211
1138	special AT-rich sequence binding protein 1 (SATB1); MAR/SAR DNA-binding protein	7.05E-08	8.35E-05	7	55	7.32982
1165	adenosine A1 receptor (ADORA1)	2.68E-07	0.000317171	7	55	19.6826
704	xeroderma pigmentosum group B complementing protein (XPB); DNA excision repair protein ERCC3; basal transcription factor 2 89-kDa subunit (BTF2-p89); TFIIH 89-kDa subunit	9.17E-07	0.00108553	7	55	3.65482
1024	trans-acting T-cell specific transcription factor GATA3	1.67E-06	0.00197407	7	55	4.28378
973	placenta growth factors 1 + 2 (PLGF1 + PLGF2)	2.02E-06	0.00239593	7	55	5.96485
929	bone morphogenetic protein 4 (BMP4) + bone morphogenetic protein 2B (BMP2B)	2.29E-06	0.00271472	7	55	8.78807
1113	putative transcription activator DB1	2.81E-06	0.00332259	7	55	6.35726
1079	janus kinase 3 (JAK3); leukocyte janus kinase (L-JAK)	3.79E-06	0.00448643	7	55	8.78321
1142	apolipoprotein E precursor (APOE)	4.29E-06	0.00507392	7	55	17.977
1100	endothelin 2 (ET2)	7.53E-06	0.00888821	7	55	10.0476
1119	high mobility group protein (HMG-I)	8.15E-06	0.00961382	7	55	6.33636
648	neogenin	8.77E-06	0.0103441	7	55	7.97533
1141	CD27L antigen receptor precursor; T-cell activation CD27 antigen	1.35E-05	0.01582	7	55	9.81295
968	bone morphogenetic protein 2A (BMP2A)	1.41E-05	0.0165598	7	55	8.78799
1177	leukocyte adhesion glycoprotein LFA-1 alpha subunit precursor; leukocyte function-associated molecule 1 alpha chain; CD11A antigen; integrin alpha L (ITGAL)	1.72E-05	0.0202037	7	55	9.13135
1073	CD40 receptor-associated factor 1 (CRAF1)	1.75E-05	0.0205453	7	55	7.31682
738	phosphatidylinositol 4-kinase alpha (PI4-kinase; PTDINS-4-kinase; PI4K-alpha)	1.92E-05	0.0225222	7	55	2.85144
1135	ras-related protein RAB-7	2.01E-05	0.0235926	7	55	7.07936
1147	guanine nucleotide-binding protein G(I)/G(S)/G(T) beta subunit 1 (GNB1); transducin beta-1 subunit	2.14E-05	0.0250426	7	55	17.1866
1152	NADH-ubiquinone oxidoreductase B18 subunit; complex I-B18 (CI-B18); cell adhesion protein SQM1	2.20E-05	0.0257559	7	55	8.98353
415	DNA-binding protein SMBP-2; glial factor-1 (GF-1)	2.31E-05	0.0269582	7	55	6.20815
1086	inhibitor of apoptosis protein1 (HIAP1; API1) + IAP homolog C; TNFR2-TRAF signaling complex protein 1; MIHC	2.79E-05	0.0324739	7	55	5.09943
1122	MAP kinase-activated protein kinase 2 (MAPKAP kinase 2; MAPKAPK-2)	3.03E-05	0.0352245	7	55	4.99514
1068	ZFM1 protein alternatively spliced product	3.21E-05	0.0373446	7	55	2.20774

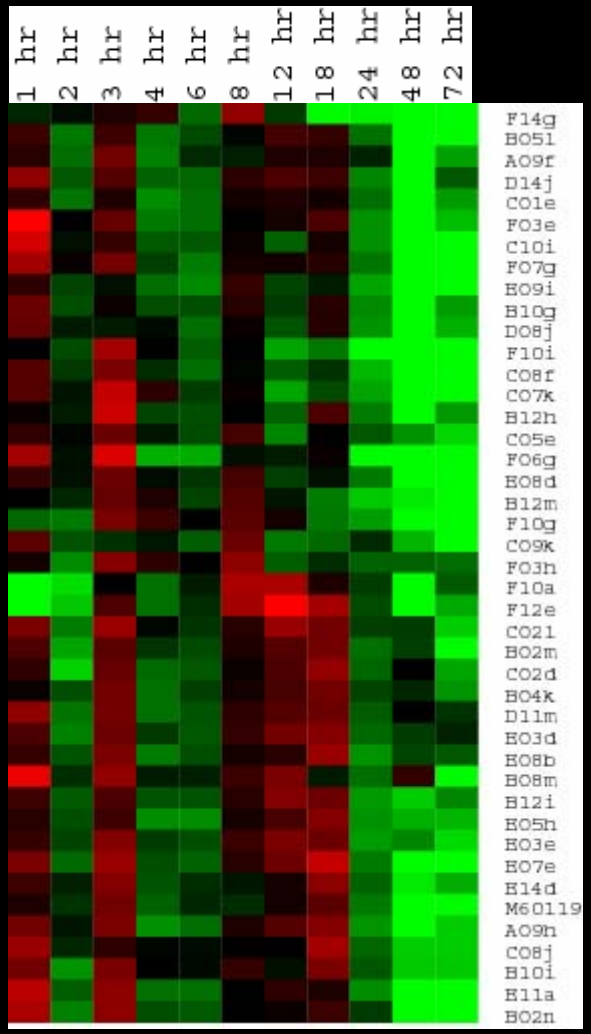
Sources of Variability in Microarray experiments

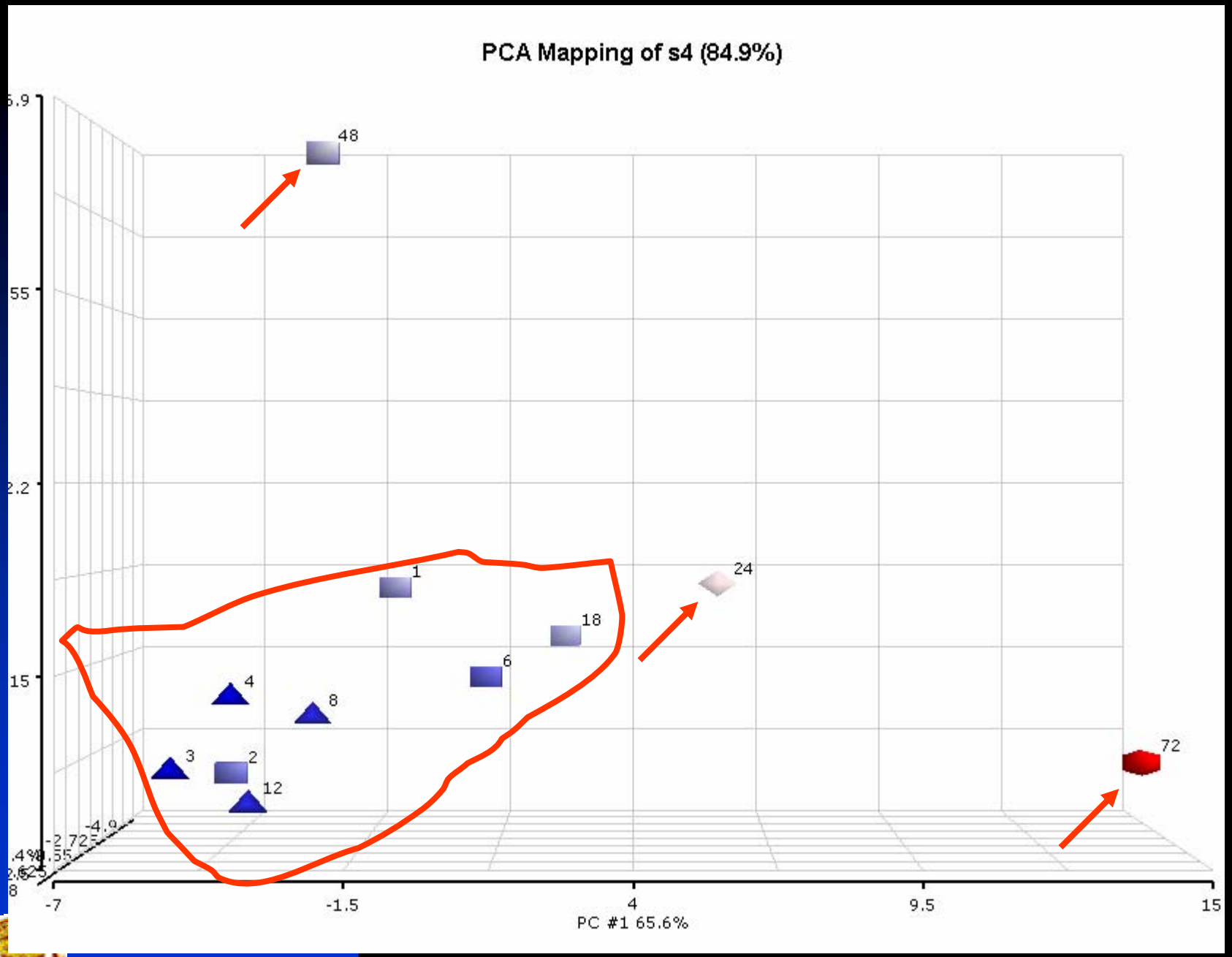
- *Biological heterogeneity in population*
- *Specimen Collection/Handling Effects;*
- *RNA extraction*
- *RNA amplification*
- *Labeling method*
- *Hybridization*
- *Scanning*
- *Image analysis*

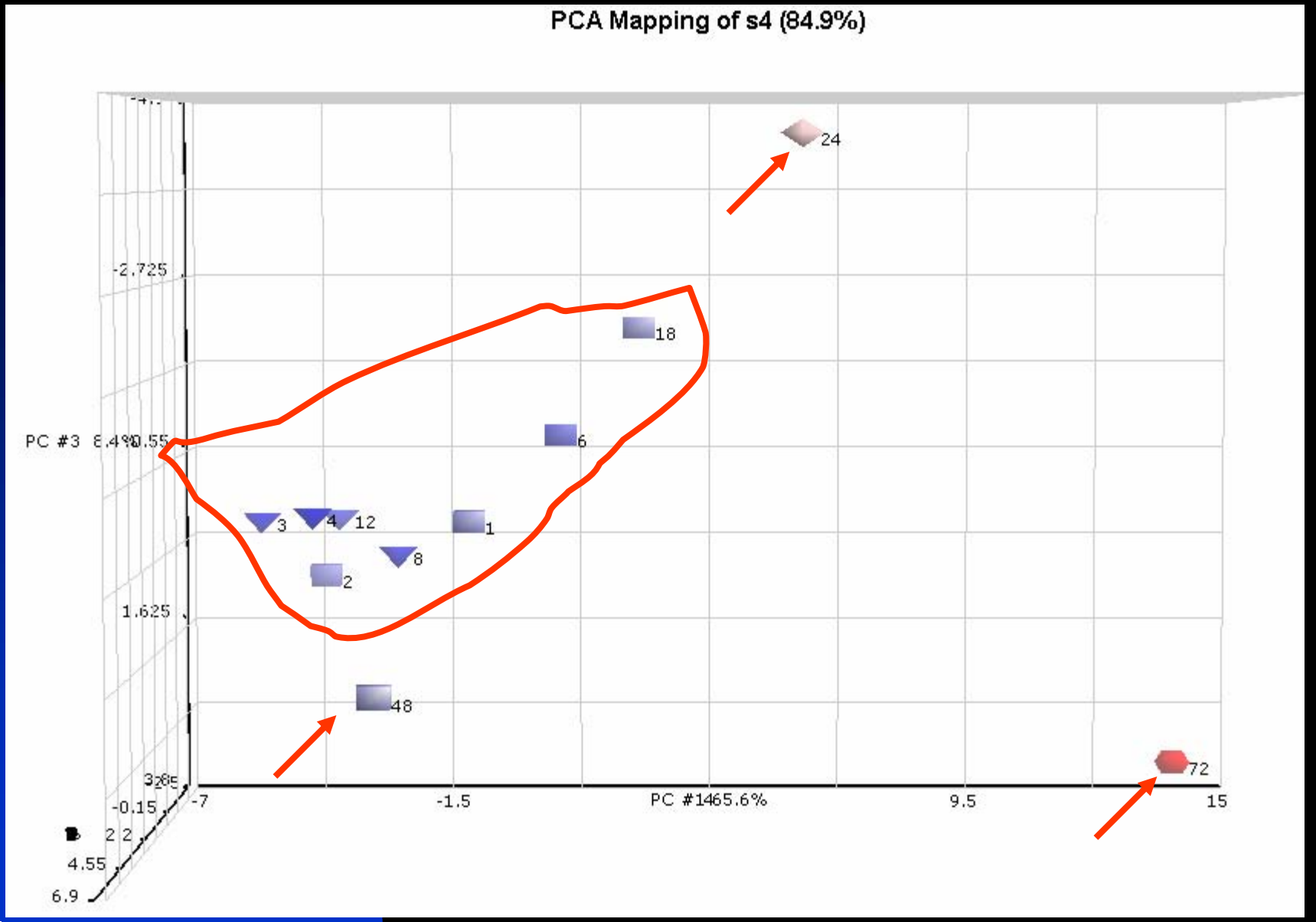
Exploring effect of time on gene expression in healthy controls

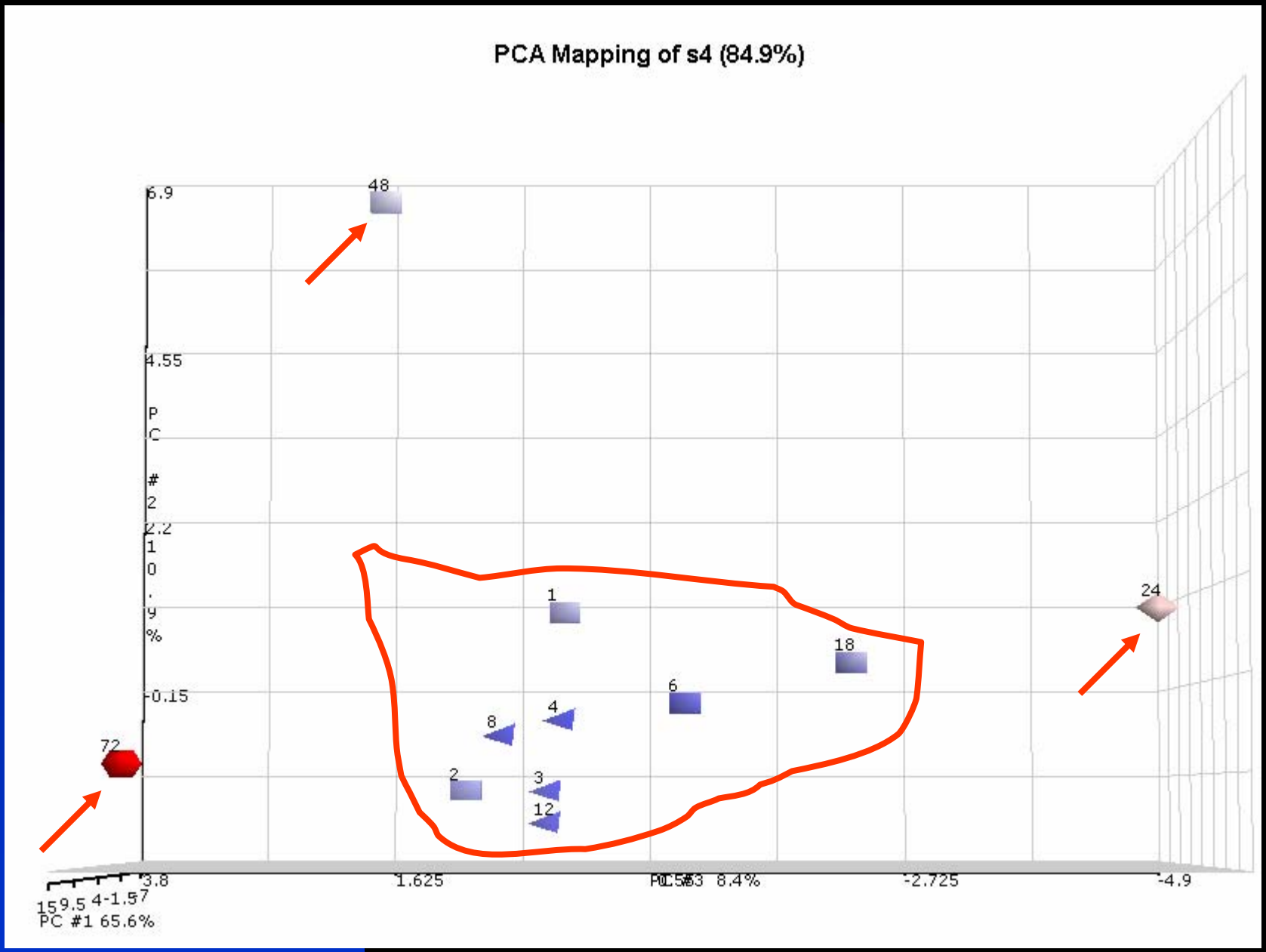
Effect of time on gene expression in the control samples






PCA Mapping of s4 (84.9%)
48
1
24
18
6
4
8
3
2
12
72
-7
-1.5
4
9.5
15
PC #1 65.6%


PCA Mapping of s4 (84.9%)
24
18
6
3
4
12
1
8
2
48
72
-2.725
1.625
-0.15
4.55
6.9
PC #3
-7
-1.5
PC #1465.6%
9.5
15


PCA Mapping of s4 (84.9%)
6.9
4.55
PC #2 2.2
10.9%
-0.15
48
72
1
18
24
6
4
8
2
3
12
3.8
1.625
-2.725
-4.9
159.5 4-1.57
PC #1 65.6%

For in vitro studies, It is critical to carry out controls tests in parallel with the treated samples in order to minimize variances in gene expression

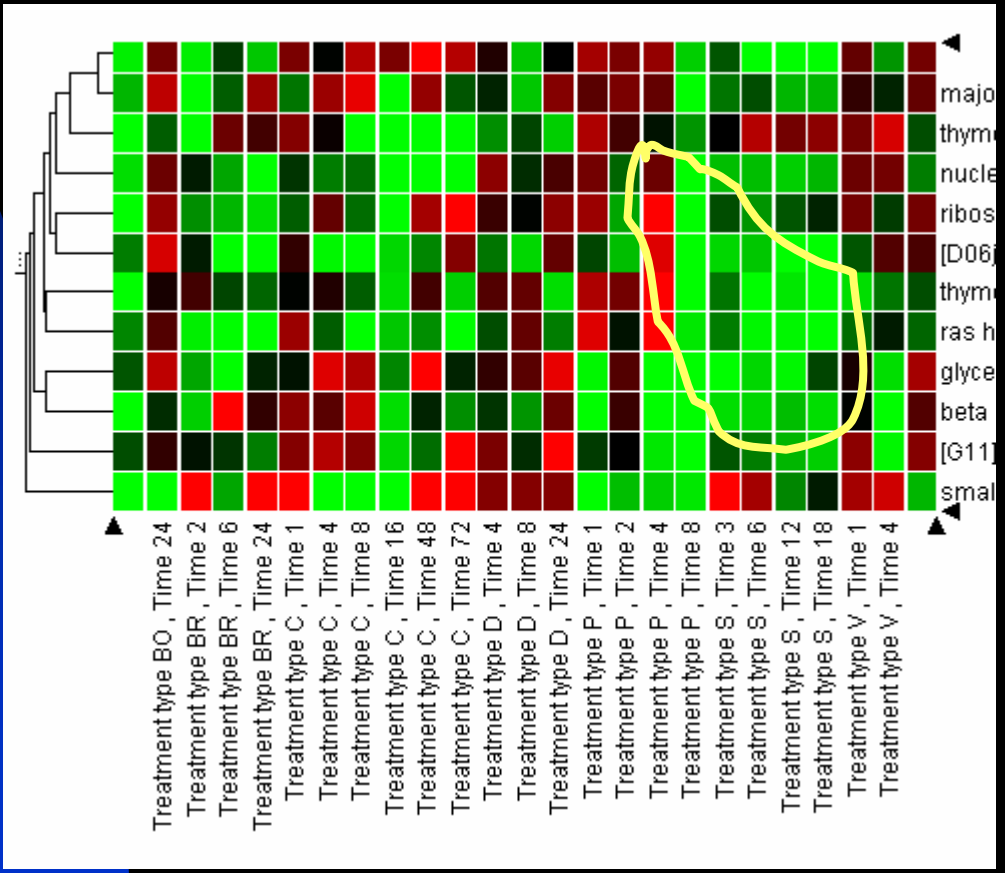
- Overview
- Establishment of base line gene expression in the host PBMC
  - Project Normal
- Data mining and feature selection
  - Genes as diagnostic markers for various pathogens.
  - *In vitro* vs. *in vivo*.
  - Applications.

- Out of these genes that were expressed below the baseline levels in all control samples and were shown to be highly expressed upon exposure to various pathogens, sets of genes were unique for certain pathogens at early time points.

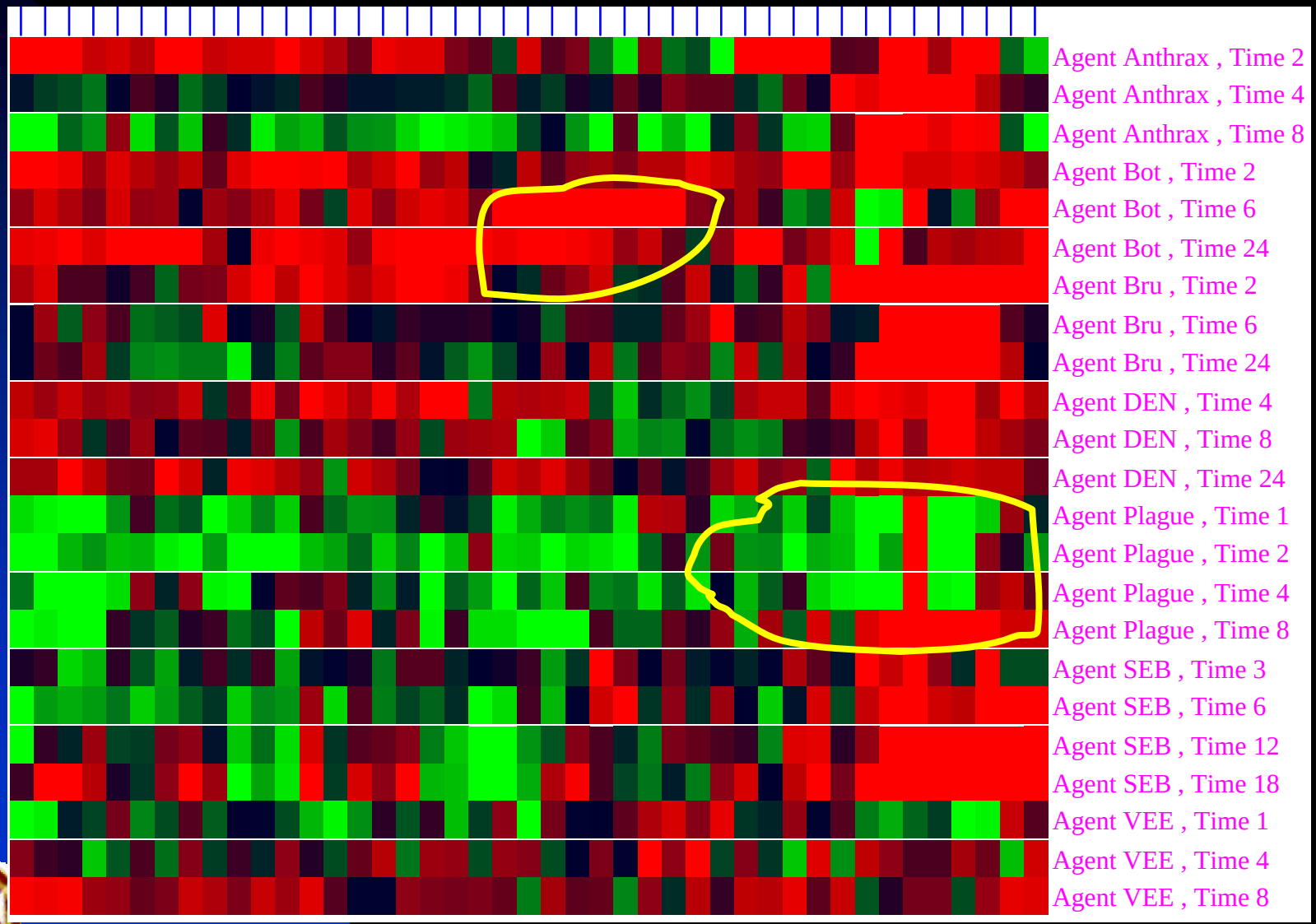
- Furthermore, genes that exhibited high expression levels in the control samples and were knocked out by exposures were also identified.

- Real time PCR was carried out to confirm gene expression patterns for some of the genes

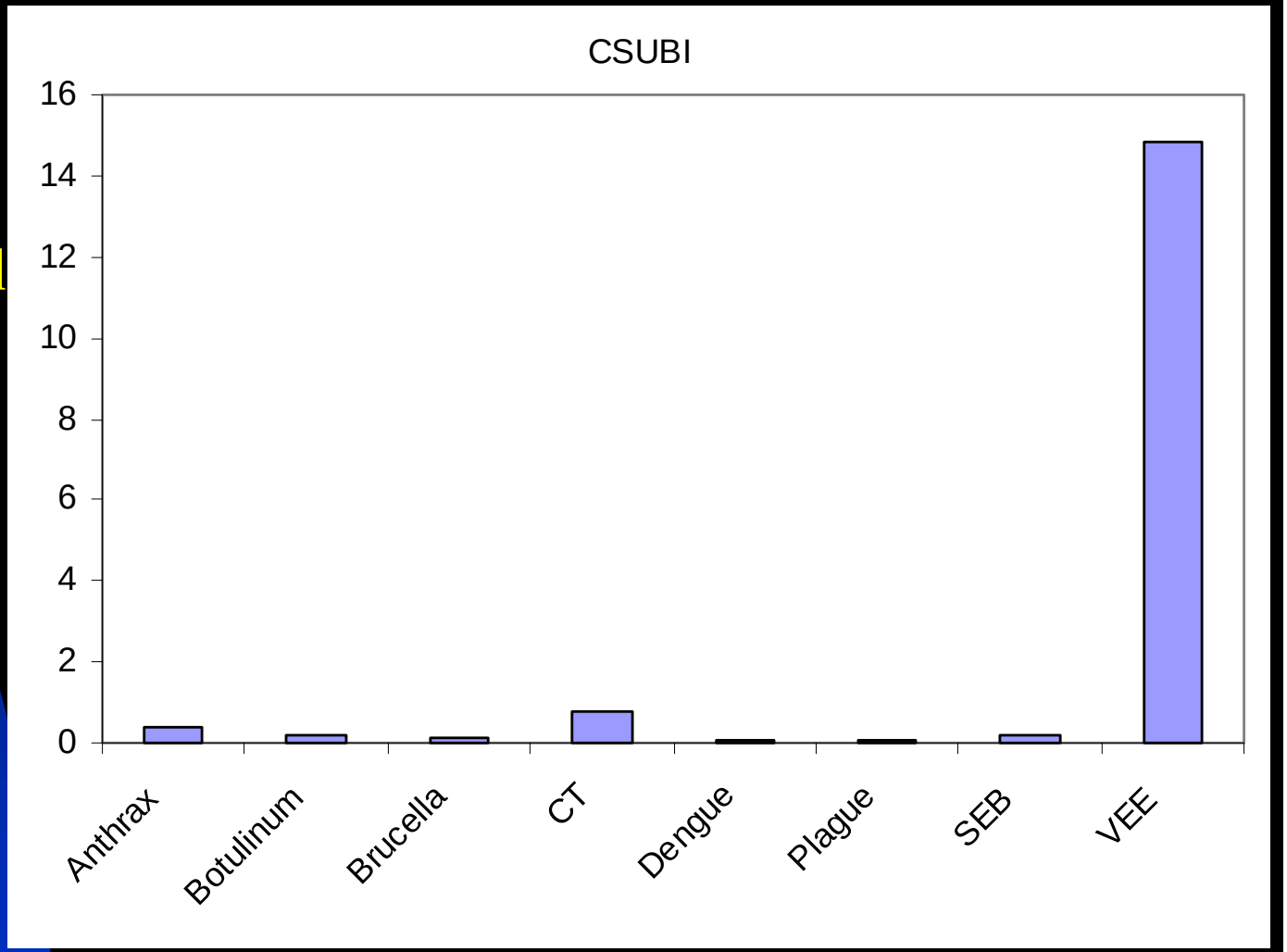
Cluster analysis of expression patterns of genes that exhibit high expression levels in the control samples

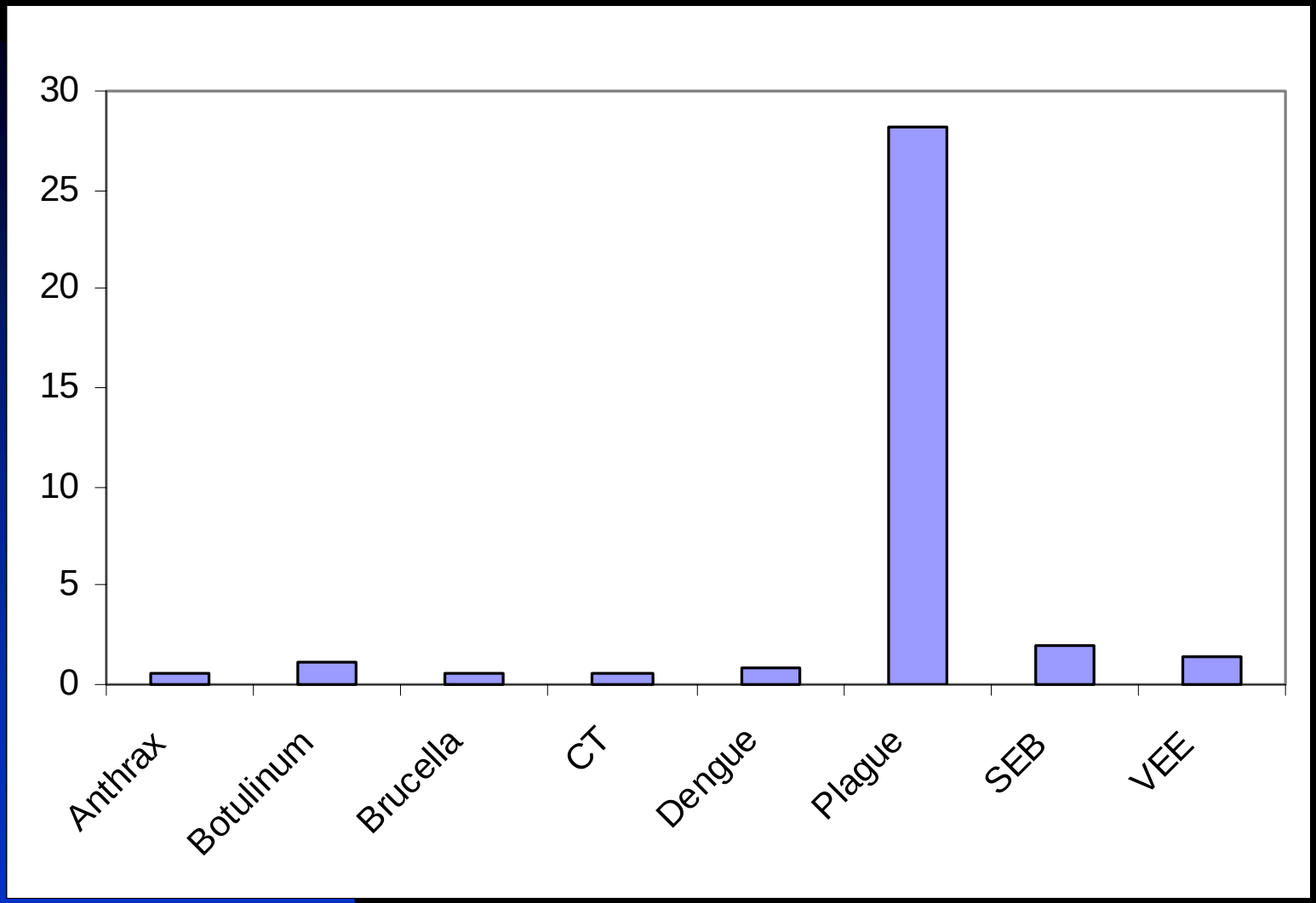


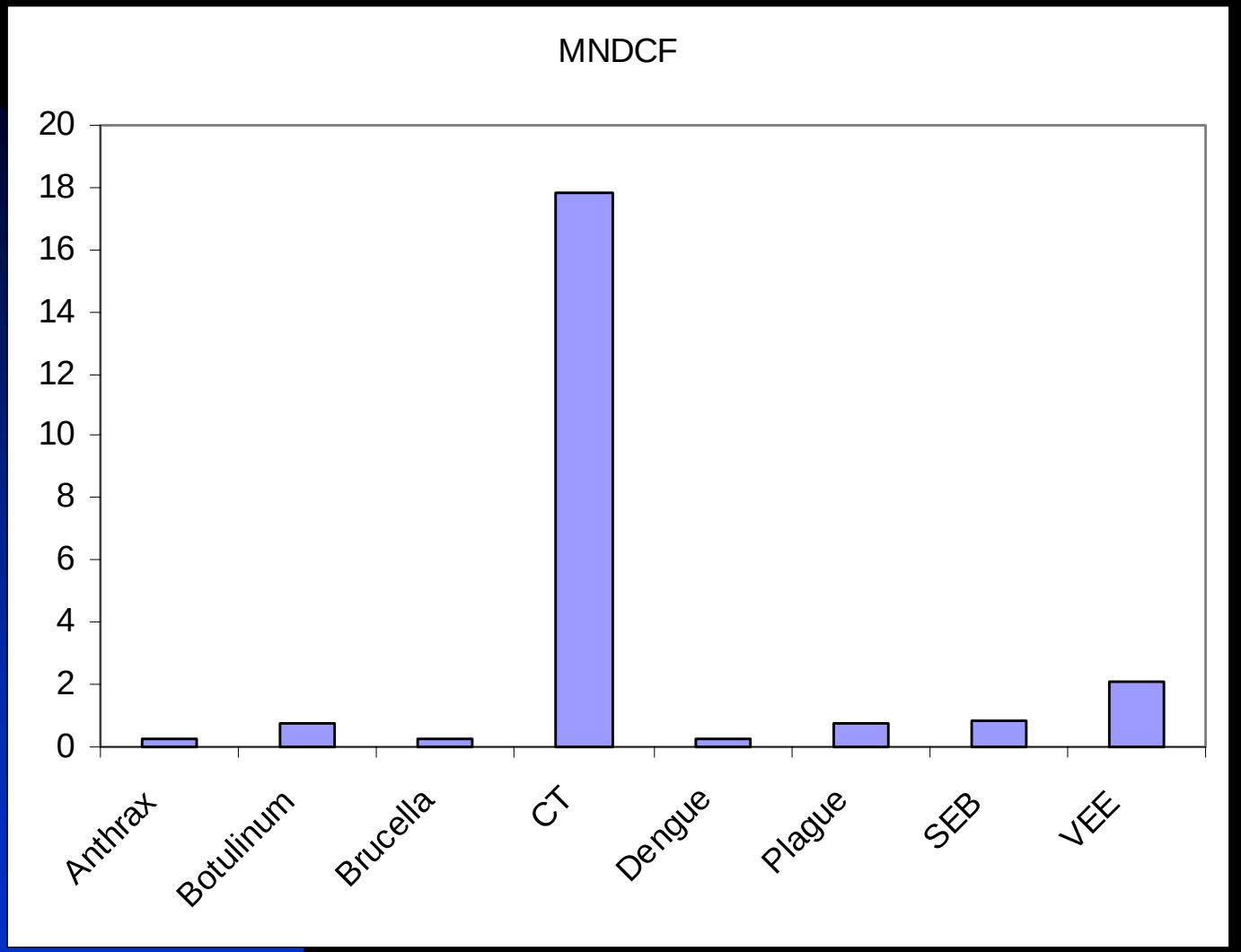
Cluster analysis of expression patterns of genes that had low expression levels in the control samples



Real time PCR for some of the genes that were expressed below background levels in the control samples and showed high expression upon exposure to one pathogen



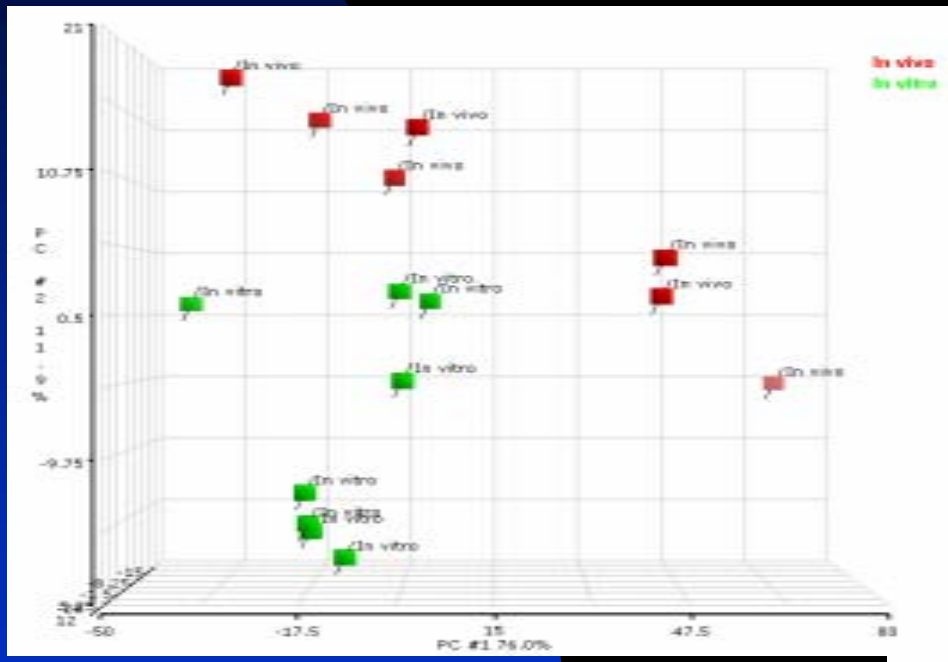

30
25
20
15
10
5
0
Anthrax
Botulinum
Brucella
CT
Dengue
Plague
SEB
VEE


MNDCF
20
18
16
14
12
10
8
6
4
2
0
Anthrax
Botulinum
Brucella
CT
Dengue
Plague
SEB
VEE

- Identify genes that exhibited on-off expression.
- Identify signature genes for some pathogens. To be continued….

In vitro vs. *in vivo*?

- Problem: Biothreat exposures are rare in humans; replicating many of these in animals is not readily possible / practical.
- Solution: Apply predictive modeling for in vitro exposures to accurately identify in vivo gene patterns.



Global gene expression profiles show a progression of SEB 2° and 3° effects w/ time

Applying predictive modeling in vitro to 8 biothreats successfully selected sets of genes to identify SEB in vivo

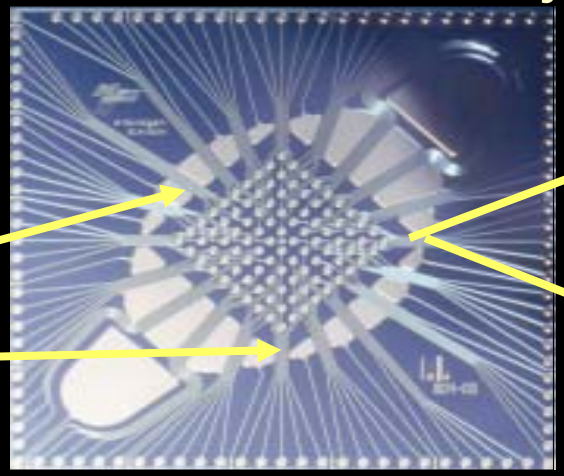
- Using *In vitro* microarray gene expression profiles, we can predict in vivo exposures using predictive modeling and cross validation algorithms.

APPLICATIONS

NANOGEN PLATFORM

Nanochip Cartridge

Electronic Microarray



Single test site

SBIR SOLICITATIONS

“Common illnesses databases”

Example: 400 genes / slide (Sept 03)

Can already carry out the complete analysis for 300 slides simultaneously.

Hand held devices for rapid diagnostics (Sept 2002)

Example: 10,000 genes analyzed within a few hours. Needs more development.

Dr. Marti Jett
Dr. Rina Das
Dr. Roger Neill
Dr. Chanaka Mendis
Dr. Sachin Mani
Dr. Christiano Cummings
Dr. Apsara Dhokalia
Dr. Mohsen Barmada
Dr. Bharati Dhruva
Dr. Maria Mayda
Nabarun Chakraborty
Shuguang Bi


QUESTION?

- The ability to identify specific gene patterns early on can provide appropriate strategies for treatment that would lead to prevention or amelioration of the disease progression.

- For some agents (*B. Anthracis*), 24 hrs is a robust response, we need to get to much earlier responses.

In vivo gene expression patterns in PBMC upon exposure to *B. Anthracis*